Berry–Esseen Bounds for Independent Sums

Contents

Contents i

1 Nonuniform Berry–Esseen bounds 1
1.1 The half-line Stein equation . . . 1
1.2 Uniform approximation and leave-one-out concentration . . . 2
1.3 Exponential concentration . . . 3
1.4 One-sided truncation . . . 4
1.5 Stein exchange and residual decomposition . . . 4
1.6 Estimates for the residual terms . . . 5
1.7 The central nonuniform estimate . . . 7
1.8 Reflection and the nonuniform theorem . . . 7

2 Normal approximation over convex sets 9
2.1 Convex distance and smooth cutoffs . . . 9
2.2 Gaussian companions and rotation . . . 11
2.3 Whitening and covariance normalization . . . 15
2.4 Signed distance and coarea . . . 16
2.5 Ball's Gaussian perimeter theorem . . . 20
2.6 Smoothing convex indicators . . . 25
2.7 Gaussian density calculus . . . 26
2.8 Taylor remainders and angle splitting . . . 28
2.9 The identity-covariance replacement induction . . . 31
2.10 Bentkus's convex-set theorem . . . 32

References 35

This volume develops two quantitative central limit theorems. Let $X_1, \dots, X_n$ be independent centered real random variables with unit total variance, and put $W = \sum_i X_i$ and $\beta_3 = \sum_i \mathbb{E}|X_i|^3$. The nonuniform Berry–Esseen theorem gives a universal constant $C > 0$ such that

$$|\Pr(W \le x) - \Phi(x)| \le \frac{C\beta_3}{1 + |x|^3}, \qquad x \in \mathbb{R}.$$

For independent centered random vectors in $\mathbb{R}^d$, put $W = \sum_i X_i$, $\Sigma = \operatorname{Cov}(W)$, and $\beta = \sum_i \mathbb{E}\|\Sigma^{-1/2} X_i\|_2^3$. If Σ is positive definite and $Z \sim \mathcal{N}_d(0, \Sigma)$, Bentkus's theorem gives a universal constant $C > 0$ such that

$$\sup_{A \in \mathcal{C}_d} |\Pr(W \in A) - \Pr(Z \in A)| \le C d^{1/4} \beta,$$

where $\mathcal{C}_d$ is the class of Borel convex subsets of $\mathbb{R}^d$. The exposition develops the proofs from their mathematical prerequisites: Stein's equation, concentration and truncation for the scalar theorem; convex distance, Gaussian boundary measure, Ball's perimeter estimate, Gaussian replacement, induction, and whitening for the multivariate theorem.

1 Nonuniform Berry–Esseen bounds

Let $\Phi(x) = \Pr(Z \le x)$ for a standard normal random variable Z. In this section I is finite, $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $(X_i)_{i\in I}$ is an independent measurable real-valued family. We write $W = \sum_i X_i$ and $\gamma = \sum_i \mathbb{E}|X_i|^3$ whenever the moments exist.

The proof proceeds from the half-line Stein equation through concentration, one-sided truncation, and an exact residual decomposition. The final reflection step converts the positive-threshold estimate into a bound on the whole real line.

1.1 The half-line Stein equation

Definition 1.1.1

Indicator Stein solution. For $z, w \in \mathbb{R}$, put

$$h_z(w) = \mathbf{1}_{\{w\le z\}} - \Phi(z), \qquad f_z(w) = e^{w^2/2}\int_{-\infty}^{w} h_z(t)e^{-t^2/2}\,dt.$$

The function f_z is continuous and nonnegative. At every $w \neq z$ it is differentiable and satisfies

$$f_z'(w) - wf_z(w) = h_z(w).$$

With the everywhere-defined representative $f_z'(w) = wf_z(w) + h_z(w)$, one has, for every $z, w \in \mathbb{R}$,

$$|wf_z(w)| \le 1, \qquad |f_z'(w)| \le 2, \qquad 0 \le f_z(w) \le \frac{\sqrt{2\pi}}{2}.$$

This is the half-line specialization of Chen and Shao (2005), Section 2.2, Lemma 2.2 and equations (2.3)–(2.9), printed pp. 10–11. The displayed integral fixes the same $(-\infty, z]$ convention as the library CDF.

Theorem 1.1.2

Tail-sensitive derivative bounds and the integrated Stein equation. Write Φ for the standard Gaussian distribution function. The derivative of the indicator Stein solution has the two exact forms

$$f_x'(w) = (1-\Phi(x))\left(1 + w\sqrt{2\pi}\, e^{w^2/2}\Phi(w)\right), \quad w \le x,$$
$$f_x'(w) = \Phi(x)\left(w\sqrt{2\pi}\, e^{w^2/2}(1-\Phi(w)) - 1\right), \quad x < w.$$

In particular, if $x > 0$ and $w \le 0$, then

$$|f_x'(w)| \le 2(1-\Phi(x)),$$

while, if $x \ge 2$ and $|w| \le x/2$, then

$$|f_x'(w)| \le (1-\Phi(x))\left(1 + \sqrt{2\pi}\,\frac{x}{2} e^{x^2/8}\right).$$

Finally, for a finite sum $W = \sum_i X_i$ of measurable square-integrable variables,

$$\Pr(W \le x) - \Phi(x) = \mathbb{E} f_x'(W) - \mathbb{E}[W f_x(W)].$$

The two closed forms are the half-line specialization of Chen and Shao (2005), equation (8.3), printed p. 54, derived from the solution in equation (2.3), printed pp. 9–10. The Mills bounds in equation (8.1) and sign estimates in equations (8.4)–(8.5), printed pp. 54–55, give the displayed tail majorants; the $|w| \le x/2$ specialization is used in the proof of equation (6.17), printed p. 46. The integrated identity is equation (1.4), printed p. 4, and the first equality in the residual expansion (6.16), printed p. 46.

1.2 Uniform approximation and leave-one-out concentration

Lemma 1.2.1

Uniform leave-one-out concentration. Suppose $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 < \infty$, and $\sum_i \operatorname{Var}(X_i) = 1$. If $\gamma < \infty$, then, for $W^{(i)} = W - X_i$, every $i \in I$, and all $a \le b$,

$$\Pr(a \le W^{(i)} \le b) \le \sqrt{2}\,(b-a) + 2(\sqrt{2}+1)\gamma.$$

The proof follows the concentration-kernel argument of Chen and Shao (2001), Section 3, especially Proposition 3.2 and equations (3.2)–(3.6), printed pp. 239–241. The constants above follow without the paper's additional small-truncation hypothesis; under that hypothesis, Chen–Shao obtain the sharper coefficients 1.5 and 3.3δ.

Theorem 1.2.2

Uniform third-moment Berry–Esseen theorem. Suppose $\mathbb{E}X_i = 0$, $\mathbb{E}|X_i|^3 < \infty$, and $\sum_i \operatorname{Var}(X_i) = 1$. Then, for every $x \in \mathbb{R}$,

$$|\Pr(W \le x) - \Phi(x)| \le 30\gamma.$$

Combining the concentration inequality with the Stein residual estimate in the manner of Chen and Shao (2005), Sections 5.1–5.2, printed pp. 31–35, gives the constant 30. The sharper constant 4.1 in Chen and Shao (2001), Theorem 2.1, printed pp. 237–238, belongs to their truncated-moment bound.

1.3 Exponential concentration

Lemma 1.3.1

Bennett–Hoeffding moment-generating-function bound. Let J be finite and let $(\eta_j)_{j\in J}$ be independent measurable random variables with finite second moments. Suppose $\mathbb{E}\eta_j \le 0$, $\eta_j \le \alpha$ almost surely, $\alpha > 0$, and

$$\sum_{j\in J} \mathbb{E}\eta_j^2 \le B^2.$$

Then, for every $t \ge 0$,

$$\mathbb{E}\exp\left(t\sum_{j\in J}\eta_j\right) \le \exp\left(\frac{e^{t\alpha}-1-t\alpha}{\alpha^2}B^2\right).$$

Consequently, for every $x \in \mathbb{R}$,

$$\Pr\left(\sum_{j\in J}\eta_j \ge x\right) \le \exp\left(-tx + \frac{e^{t\alpha}-1-t\alpha}{\alpha^2}B^2\right).$$

This is Lemma 6.2 and equation (6.2) of Chen and Shao (2005), together with its Chernoff consequence, printed pp. 40–41. Keeping t as a parameter and optimizing it recovers the paper's equations (6.3)–(6.4).

Proposition 1.3.2

Exponential concentration for the upper-truncated leave-one-out sum. Define $\bar{X}_j = X_j \mathbf{1}_{\{X_j \le 1\}}$ and $\bar{W}^{(i)} = \sum_{j\ne i}\bar{X}_j$. Suppose $\mathbb{E}X_j = 0$, $\mathbb{E}X_j^2 < \infty$, $\sum_j \operatorname{Var}(X_j) = 1$, and $\gamma < \infty$. Then, for every $i \in I$ and all $a \le b$,

$$\Pr(a \le \bar{W}^{(i)} \le b) \le e^{-a/2}\big(24(b-a) + 48\gamma\big).$$

This is the exponentially weighted conclusion of Chen and Shao (2005), Proposition 6.1, equation (6.1), with the proof on printed pp. 41–43. The paper prints constants $5, 7$. Accounting explicitly for the one-sided truncation drift gives the coefficients $24, 48$; the displayed exchange equality in the source omits this drift.

1.4 One-sided truncation

Proposition 1.4.1

Comparison with one-sided truncation. Suppose $\mathbb{E}X_i = 0$, $\mathbb{E}|X_i|^3 < \infty$, $\sum_i \operatorname{Var}(X_i) = 1$, and $\gamma \le 1$. Put $\bar{X}_i = X_i \mathbf{1}_{\{X_i \le 1\}}$ and $\bar{W} = \sum_i \bar{X}_i$. For every $z \ge 2$,

$$\Pr(W > z) \le \Pr(\bar{W} > z) + \frac{45\gamma}{1+z^3},$$

and hence

$$|\Pr(W \le z) - \Phi(z)| \le |\Pr(\bar{W} \le z) - \Phi(z)| + \frac{45\gamma}{1+z^3}.$$

The event decomposition is in Section 6.2, equations (6.13)–(6.14), of Chen and Shao (2005), printed p. 44. The displayed 45 is the explicit constant proved by the large-jump and leave-one-out estimates.

1.5 Stein exchange and residual decomposition

Theorem 1.5.1

Noncentered Stein exchange for the upper-truncated sum. Let $Y_i = \bar{X}_i$, $\bar{W} = \sum_i Y_i$, and $\bar{W}^{(i)} = \bar{W} - Y_i$. Define

$$K_i(t) = \mathbb{E}\left[|Y_i| \mathbf{1}_{\{\min(0,Y_i) \le t \le \max(0,Y_i)\}}\right].$$

If the X_i have finite second moments, then, for every $z \in \mathbb{R}$,

$$\mathbb{E}[\bar{W} f_z(\bar{W})] = \sum_i \int_{\mathbb{R}} \mathbb{E}[f_z'(\bar{W}^{(i)} + t)] K_i(t)\, dt + \sum_i \mathbb{E}Y_i \, \mathbb{E}[f_z(\bar{W}^{(i)})].$$

This is the noncentered form of the exchange used in Section 6.2, equation (6.16), of Chen and Shao (2005), printed p. 45. Retaining the second sum is essential: one-sided truncation generally makes $\mathbb{E}Y_i < 0$.

Theorem 1.5.2

The $R_1 + R_2 + R_3$ identity. Set

$$q = 1 - \sum_i \mathbb{E}\bar{X}_i^2, \qquad T = \sum_i \int_{\mathbb{R}} \mathbb{E}[f_z'(\bar{W}^{(i)} + t)] K_i(t)\, dt,$$

and

$$M = \sum_i \mathbb{E}\bar{X}_i \, \mathbb{E}[f_z(\bar{W}^{(i)})].$$

Define

$$R_1 = q\,\mathbb{E}f_z'(\bar{W}), \qquad R_2 = (1-q)\mathbb{E}f_z'(\bar{W}) - T, \qquad R_3 = -M.$$

If $\mathbb{E}X_i = 0$, $\sum_i \mathrm{Var}(X_i) = 1$, and $\gamma < \infty$, then $0 \le q \le \gamma$; moreover, without the centering and variance assumptions needed for that bound, the exact identity

$$\Pr(\bar{W} \le z) - \Phi(z) = R_1 + R_2 + R_3$$

holds whenever the coordinates have finite second moments.

This is Section 6.2, equation (6.16), of Chen and Shao (2005), printed p. 45, with the missing second-moment mass and nonzero truncated means made explicit.

1.6 Estimates for the residual terms

Lemma 1.6.1

Exponential bounds for R_1 and R_3. Under the centered, unit-total-variance, finite-third-moment hypotheses above, for every $z \ge 2$,

$$|R_1| \le 8e^{-z/2}\gamma, \qquad |R_3| \le 8e^{-z/2}\gamma.$$

These are explicit versions of Chen and Shao (2005), Section 6.2, equations (6.17)–(6.18), printed p. 45. The source records an unspecified absolute constant; the estimates above give 8 in both inequalities.

Theorem 1.6.2

Deterministic expected-kernel representation and split of R_2. For $K_i(t) = \mathbb{E}[|\bar{X}_i|\mathbf{1}_{\{\min(0,\bar{X}_i) \le t \le \max(0,\bar{X}_i)\}}]$, one has

$$\int_{\mathbb{R}} K_i(t)\,dt = \mathbb{E}\bar{X}_i^2.$$

Writing $F = \mathbb{E}f_z'(\bar{W})$, the residual is

$$R_2 = \sum_i \int_{\mathbb{R}} \big(F - \mathbb{E}f_z'(\bar{W}^{(i)} + t)\big) K_i(t)\,dt.$$

Using $f_z'(w) = w f_z(w) + \mathbf{1}_{\{w \le z\}} - \Phi(z)$ splits this exactly as $R_2 = R_{2,1} + R_{2,2}$, where

$$R_{2,1} = \sum_i \int \big(\Pr(\bar{W} \le z) - \Pr(\bar{W}^{(i)} + t \le z)\big) K_i(t)\,dt$$

and

$$R_{2,2} = \sum_i \int \big(\mathbb{E}[\bar{W} f_z(\bar{W})] - \mathbb{E}[(\bar{W}^{(i)} + t) f_z(\bar{W}^{(i)} + t)]\big) K_i(t)\,dt.$$

This is the split immediately after equation (6.18) in Chen and Shao (2005), Section 6.2, printed pp. 45–46. Fubini and leave-one-out independence supply the deterministic kernels in the display.

Lemma 1.6.3

> **Indicator part of R_2.** Under the centered, unit-total-variance, finite-third-moment hypotheses, for every $z \in \mathbb{R}$,
>
> $$|R_{2,1}| \le 168 e^{-z/2} \gamma.$$

Applying the two orientations of the conditional interval estimate in Section 6.2, equations (6.19) and (6.21)–(6.22), of Chen and Shao (2005), printed p. 46 gives the stated bound. The source uses an unspecified constant; the preceding concentration estimate gives 168.

Lemma 1.6.4

> **Stein-product increment estimate.** Let $g_z(w) = (w f_z(w))'$ away from z, with the everywhere-defined representative obtained from the two one-sided formulas. There is an explicit absolute constant $C_{\mathrm{inc}} > 0$ such that, under the centered and unit-total-variance hypotheses, for every $i \in I$, $z \ge 2$, and $s, t \le 1$,
>
> $$\left| \mathbb{E}\big[(\bar{W}^{(i)} + t) f_z(\bar{W}^{(i)} + t) - (\bar{W}^{(i)} + s) f_z(\bar{W}^{(i)} + s) \big] \right| \le C_{\mathrm{inc}} e^{-z/2} (|s| + |t|).$$
>
> In addition,
>
> $$b f_z(b) - a f_z(a) = \int_a^b g_z(u)\, du$$
>
> for every $a, b \in \mathbb{R}$, including intervals crossing z.

This is Lemma 6.5 and equations (6.23)–(6.24) of Chen and Shao (2005), printed pp. 46–48. Here $C_{\mathrm{inc}} = C_{\mathrm{low}} + 3(\sqrt{2\pi}/2 + 2) e^2 e^{e^2 - 3}$.

Lemma 1.6.5

> **Stein-product part of R_2.** Leave-one-out independence gives, for every i,
>
> $$\mathbb{E}[\bar{W} f_z(\bar{W})] = \mathbb{E}_{\bar{X}_i} \mathbb{E}_{\bar{W}^{(i)}} [(\bar{W}^{(i)} + \bar{X}_i) f_z(\bar{W}^{(i)} + \bar{X}_i)].$$
>
> Consequently, under the centered, unit-total-variance, finite-third-moment hypotheses and for every $z \ge 2$,
>
> $$|R_{2,2}| \le \frac{3}{2} C_{\mathrm{inc}} e^{-z/2} \gamma.$$

This completes the integration step in Chen and Shao (2005), Section 6.2, equation (6.20), printed p. 46. The factor $3/2$ is the sum of the exact first absolute moment of the expected kernel, $1/2$, and the product of its total mass with the coordinate first moment, bounded by 1.

1.7 The central nonuniform estimate

Theorem 1.7.1

Central exponential estimate for the upper-truncated sum. There is an absolute constant

$$A = 184 + \frac{3}{2} C_{\mathrm{inc}} \geq 0$$

such that, for every finite independent measurable family with $\mathbb{E}X_i = 0$, $\mathbb{E}|X_i|^3 < \infty$, and $\sum_i \mathrm{Var}(X_i) = 1$, for every $z \geq 2$,

$$\left|\Pr(\bar{W} \leq z) - \Phi(z)\right| \leq A e^{-z/2} \gamma.$$

This is the finite-third-moment specialization of the conclusion assembled in Section 6.2, equations (6.16)–(6.24), of Chen and Shao (2005), printed pp. 45–48. The four explicit contributions are 8 from R_1, 8 from R_3, 168 from $R_{2,1}$, and $(3/2)C_{\mathrm{inc}}$ from $R_{2,2}$.

1.8 Reflection and the nonuniform theorem

Theorem 1.8.1

Reduction from upper-truncated decay to the nonuniform theorem. Suppose an absolute $A \geq 0$ satisfies the central exponential estimate in the preceding node for all finite families. Then there is an absolute $C = 270 + 54A > 0$ such that every centered independent family with finite third moments and unit total variance satisfies, for all $x \in \mathbb{R}$,

$$|\Pr(W \leq x) - \Phi(x)| \leq \frac{C\gamma}{1 + |x|^3}.$$

The positive-threshold argument follows Chen and Shao (2005), Section 6.2, printed pp. 44–48: the uniform theorem handles bounded x, the truncation comparison handles large jumps, and $e^{-x/2} \leq 54/(1 + x^3)$ for $x \geq 2$. Reflection of the law supplies negative thresholds with the atom-safe $(-x) + \varepsilon$ limit required by the closed-half-line CDF convention.

Theorem 1.8.2

Finite-third-moment nonuniform Berry–Esseen theorem. There is an absolute constant $C > 0$ with the following property. Let I be a finite index set, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(X_i)_{i \in I}$ be independent measurable real-valued random variables satisfying

$$\mathbb{E}X_i = 0, \qquad \mathbb{E}|X_i|^3 < \infty, \qquad \sum_{i \in I} \mathrm{Var}(X_i) = 1.$$

Writing

$$W = \sum_{i \in I} X_i, \qquad \gamma = \sum_{i \in I} \mathbb{E}|X_i|^3,$$

for every $x \in \mathbb{R}$ one has

$$|\Pr(W \le x) - \Phi(x)| \le \frac{C\gamma}{1 + |x|^3},$$

where Φ is the distribution function of a standard normal random variable.

This finite-third-moment form goes back to Bikelis (1966). Theorem 2.2 and the discussion following it in Chen and Shao (2001), printed pp. 238–239, prove the stronger truncated-moment version. The proof presented here follows Section 6.2 of Chen and Shao (2005), printed pp. 44–48, including the universal $R_{2,2}$ estimate and an atom-safe reflection argument.

2 Normal approximation over convex sets

Throughout this section $E = \mathbb{R}^d$ has its Euclidean norm. For $A \subseteq E$, write

$$d_A(x) = \inf_{a \in A} \|x - a\|, \qquad A^\varepsilon = \{x : d_A(x) < \varepsilon\}.$$

The proof has four layers. Convex distance produces smooth indicator approximations; signed distance and coarea reduce Gaussian shells to boundary measure; Ball's projection argument bounds that boundary measure by the sharp order $d^{1/4}$; and a Gaussian replacement induction, followed by covariance whitening, proves the multivariate limit theorem.

2.1 Convex distance and smooth cutoffs

Lemma 2.1.1

> **Convex parallel sets.** If $A \subseteq \mathbb{R}^d$ is convex, then its open parallel set A^ε is convex for every $\varepsilon \in \mathbb{R}$; its closed parallel set $\{x : d_A(x) \le \varepsilon\}$ is convex whenever $\varepsilon \ge 0$. Both parallel sets are measurable, and for every $v \in \mathbb{R}^d$,
>
> $$(v + A)^\varepsilon = v + A^\varepsilon$$
>
> with the analogous identity for closed parallel sets.

These are the set-theoretic facts used in Bentkus (2004), condition (ii) and the neighborhood notation preceding Theorem 1.2, printed pp. 401–402. The open-thickening identity remains valid for all real radii.

Theorem 2.1.2

> **Metric projection onto a closed convex set.** Let $A \subseteq \mathbb{R}^d$ be nonempty, closed, and convex. For every x there is a unique point $p_A(x) \in A$ such that $\|x - p_A(x)\| = d_A(x)$. It satisfies
>
> $$\langle x - p_A(x), a - p_A(x) \rangle \le 0 \quad (a \in A)$$

and, for all x, y,

$$\|p_A(x) - p_A(y)\|^2 \le \langle x - y, p_A(x) - p_A(y)\rangle.$$

Consequently both p_A and the residual map $x \mapsto x - p_A(x)$ are 1-Lipschitz.

This packages the Euclidean projection geometry used in Bentkus (2003), Lemma 2.2, printed pp. 389–390. The paper states the resulting distance-gradient inequalities; the projection formulation records the geometric mechanism behind them.

Theorem 2.1.3

Differentiability of distance to a convex set. Under the hypotheses of the preceding node, for every $x, h \in \mathbb{R}^d$,

$$\left|d_A(x+h)^2 - d_A(x)^2 - 2\langle x - p_A(x), h\rangle\right| \le 3\|h\|^2.$$

Thus d_A^2 is Fréchet differentiable everywhere with

$$D(d_A^2)(x)[h] = 2\langle x - p_A(x), h\rangle.$$

If $x \notin A$, then d_A is Fréchet differentiable at x,

$$Dd_A(x)[h] = \left\langle \frac{x - p_A(x)}{d_A(x)}, h \right\rangle, \qquad \|Dd_A(x)\| = 1.$$

This is the distance calculus behind Lemma 2.2, equations (2.5)–(2.9), of Bentkus (2003), printed pp. 389–390. The derivative is written as a bounded linear functional, together with an explicit quadratic remainder.

Theorem 2.1.4

Bentkus's continuously differentiable distance cutoff. Let $A \subseteq \mathbb{R}^d$ be nonempty, closed, and convex, and let $\varepsilon > 0$. There is a continuously differentiable function $\varphi_{\varepsilon,A} : \mathbb{R}^d \to [0,1]$ of the form

$$\varphi_{\varepsilon,A}(x) = \psi(d_A(x)/\varepsilon)$$

such that

$$\varphi_{\varepsilon,A} = 1 \text{ on } A, \qquad \varphi_{\varepsilon,A} = 0 \text{ when } d_A(x) \ge \varepsilon,$$

and, for every $x, y \in \mathbb{R}^d$,

$$\|D\varphi_{\varepsilon,A}(x)\| \le \frac{2}{\varepsilon}, \qquad \|D\varphi_{\varepsilon,A}(x) - D\varphi_{\varepsilon,A}(y)\| \le \frac{8}{\varepsilon^2}\|x - y\|.$$

This is Lemma 2.3, equations (2.10)–(2.12), of Bentkus (2003), printed pp. 390–391, and the same cutoff imported in Bentkus (2004), Lemma 2.2, equation (2.2), printed p. 402. Here ψ is the paper's piecewise quadratic profile, including its endpoint values.

2.2 Gaussian companions and rotation

Theorem 2.2.1

Gaussian companions and canonical probability-space transport. Let $X_i : \Omega \to \mathbb{R}^d$ be a finite measurable family on a probability space (Ω, μ), put $\nu_i = \mathcal{L}_\mu(X_i)$, and let Σ_i be the matrix of the covariance bilinear form of ν_i. Then Σ_i is positive semidefinite, and the canonical Gaussian product

$$\Gamma = \bigotimes_i N(0, \Sigma_i)$$

has independent coordinate variables Y_i that are centered, belong to L^3, and satisfy

$$\mathrm{Cov}(Y_i)[u, v] = \mathrm{Cov}_{\nu_i}[u, v] \qquad (u, v \in \mathbb{R}^d).$$

The marginal-law product $\nu = \bigotimes_i \nu_i$ is a probability measure whose coordinate variables $\widetilde{X}_i$ are independent. If the original X_i are independent, then

$$\mathcal{L}_\mu((X_i)_i) = \nu, \qquad \mathcal{L}_\mu\left(\sum_i X_i\right) = \mathcal{L}_\nu\left(\sum_i \widetilde{X}_i\right).$$

Finally, the product $\nu \otimes \Gamma$ carries the two coordinate families $(\widetilde{X}_i)_i$ and $(Y_i)_i$ on one probability space. Each family is mutually independent, the two coordinate blocks are independent, and hence every $\widetilde{X}_i$ is independent of every Y_j. On this product space the original coordinate retains law ν_i, the Gaussian coordinate retains law $N(0, \Sigma_i)$, and their covariance bilinear forms agree.

This is the canonical-space realization of the Gaussian companions Y_i, the leave-one-out variables in equation (3.2), and the joint enlargement implicit in the telescoping/rotation setup (3.2)–(3.5) of Bentkus (2004), printed pp. 403–404. The source assumes all variables exist independently on one probability space. The canonical product construction above realizes this arrangement while allowing singular individual covariance matrices. The rotation identity (3.4) is treated below.

Lemma 2.2.2

Dimension-free fourth-moment control for a Gaussian vector. Let S be a positive semidefinite $d \times d$ real matrix and let $Y \sim N(0, S)$. Then

$$\mathbb{E}\|Y\|^4 \le 3\big(\mathbb{E}\|Y\|^2\big)^2.$$

No nonsingularity assumption is imposed on S.

This is an explicit sufficient estimate for the dimension-free Gaussian moment comparison used by Bentkus (2004), in the discussion following equation (3.35), printed p. 408. Bentkus records the resulting third-moment comparison with an implicit absolute constant; the displayed fourth-moment bound, together with Hölder's inequality, gives an explicit constant.

Lemma 2.2.3

> **Exact second-moment matching on the replacement space.** Let X_i be measurable, centered, square-integrable random vectors in $\mathbb{R}^d$. On the canonical replacement space, let $\widetilde{X}_i$ have the law of X_i and let $Y_i \sim N(0, \mathrm{Cov}(X_i))$ be its independent Gaussian companion. Then
>
> $$\mathbb{E}\|Y_i\|^2 = \mathbb{E}\|\widetilde{X}_i\|^2.$$
>
> The covariance, and hence the Gaussian law, may be singular.

This is the norm-moment consequence of the covariance matching used by Bentkus (2004), immediately after equation (3.35), printed p. 408, where the equality $\mathbb{E}|Y|^2 = \mathbb{E}|X|^2$ is invoked explicitly.

Theorem 2.2.4

> **Third-moment comparison for a covariance-matched Gaussian.** Let X be a centered L^3 random vector in $\mathbb{R}^d$, and let $Y \sim N(0, \mathrm{Cov}(X))$. With the absolute constant $C_G = 27$, one has
>
> $$\mathbb{E}\|Y\|^3 \le C_G\, \mathbb{E}\|X\|^3.$$
>
> Equivalently, for each coordinate pair $(\widetilde{X}_i, Y_i)$ on the canonical replacement space,
>
> $$\mathbb{E}\|Y_i\|^3 \le 27\, \mathbb{E}\|\widetilde{X}_i\|^3.$$

Bentkus (2004), uses $\mathbb{E}|Y|^3 \ll \mathbb{E}|X|^3$ after equation (3.35), printed p. 408, and reuses the same absolute-moment control in equations (3.39)–(3.40), printed p. 409. The conservative value 27 is one explicit choice for the source's implicit constant.

Lemma 2.2.5

> **Mixed original–Gaussian moment bound.** Under the preceding centered L^3 hypotheses, the independent canonical pair $(\widetilde{X}_i, Y_i)$ satisfies
>
> $$\mathbb{E}\big[\|\widetilde{X}_i\|^2 \|Y_i\|\big] \le \mathbb{E}\|\widetilde{X}_i\|^3.$$

This is the Hölder estimate used immediately before equation (3.30) in Vidmantas Bentkus (2004), printed p. 407. The source writes the estimate with X and its covariance-matched Gaussian Y; the display above uses their canonical independent coupling.

Theorem 2.2.6

Rotation and low-order moment cancellation. Let X, Y be independent random vectors in a real Banach space E, assume that they are integrable and centered, and let $B : E \times E \to \mathbb{R}$ be a continuous bilinear form for which $B(X, X)$ and $B(Y, Y)$ are integrable. If

$$\mathbb{E}B(X, X) = \mathbb{E}B(Y, Y),$$

then, for every $\alpha \in \mathbb{R}$, the rotation and its angular derivative

$$X_\alpha = \cos\alpha\, X + \sin\alpha\, Y, \qquad X'_\alpha = -\sin\alpha\, X + \cos\alpha\, Y$$

are centered and satisfy the quadratic cancellation

$$\mathbb{E}B(X_\alpha, X'_\alpha) = 0.$$

Pointwise, the bilinear term expands as

$$B(X_\alpha, X'_\alpha) = -\cos\alpha\sin\alpha\, B(X, X) + \cos^2\alpha\, B(X, Y) - \sin^2\alpha\, B(Y, X) + \sin\alpha\cos\alpha\, B(Y, Y).$$

In particular, for every centered measurable L^3 summand X_i, the canonical pair consisting of its marginal-law coordinate and its covariance-matched Gaussian companion satisfies this cancellation for every continuous bilinear B.

This is the low-order rotation layer introduced immediately before equation (3.4) in Bentkus (2004), and the linear/bilinear cancellation asserted in equation (3.5), printed p. 404. Independence kills the two mixed expectations and the matched second moment kills the remaining pair. The full telescoping identity (3.4) and its third-order remainder estimates are assembled in the rotation and standardized-induction nodes below.

Theorem 2.2.7

Coordinate rotation integrands and justified exchange of expectations. On the canonical replacement space $(\widetilde{X}_i, Y_i)_{i\in I}$, put

$$Z_\alpha = \sum_i (\cos\alpha\, \widetilde{X}_i + \sin\alpha\, Y_i), \qquad Z'_{k,\alpha} = -\sin\alpha\, \widetilde{X}_k + \cos\alpha\, Y_k,$$

and let $V_k = \sum_{i\neq k} Y_i$. For a convex set $A \subseteq \mathbb{R}^d$, $\varepsilon > 0$, and the Bentkus cutoff $\varphi = \varphi_{\varepsilon,A}$, define

$$T_k(\alpha) = D\varphi(Z_\alpha)[Z'_{k,\alpha}], \qquad R_k(\alpha) = D\varphi\Big(V_k + \cos\alpha\, \widetilde{X}_k + \sin\alpha\, Y_k\Big)\,[Z'_{k,\alpha}].$$

The rotated sum splits pointwise into its leave-one-out base and exposed coordinate:

$$Z_\alpha = \cos\alpha \sum_{i\neq k} \widetilde{X}_i + \sin\alpha \sum_{i\neq k} Y_i + \cos\alpha\, \widetilde{X}_k + \sin\alpha\, Y_k.$$

If the original coordinates are measurable and belong to L^3, then on every bounded angle interval both coordinate integrands are jointly integrable with respect to angle and the replacement law. Consequently,

$$\int_a^b \mathbb{E}T_k(\alpha)\,d\alpha = \mathbb{E}\int_a^b T_k(\alpha)\,d\alpha, \qquad \int_a^b \mathbb{E}R_k(\alpha)\,d\alpha = \mathbb{E}\int_a^b R_k(\alpha)\,d\alpha.$$

If $a \le b$, finite summation may also be interchanged:

$$\mathbb{E}\left[\sum_k \int_a^b T_k(\alpha)\,d\alpha\right] = \sum_k \int_a^b \mathbb{E}T_k(\alpha)\,d\alpha.$$

Bentkus (2004), introduces the uniform rotation and leave-one-out sums immediately before equation (3.4), printed pp. 403–404; equations (3.4)–(3.5), printed p. 404, give the coordinate representation and low-order cancellations. The actual and fully Gaussian reference contributions are the terms split in equations (3.11)–(3.12), also printed p. 404. The joint integrability assertions above justify the Fubini interchanges implicit in the paper's compact notation.

Theorem 2.2.8

Covariance additivity and the leave-one-out identity. Let $X_1, \dots, X_n$ be independent square-integrable random vectors in $\mathbb{R}^d$. For every set of indices J and every $u, v \in \mathbb{R}^d$, one has

$$\mathrm{Cov}\left(\sum_{i\in J} X_i\right)[u,v] = \sum_{i\in J} \mathrm{Cov}(X_i)[u,v].$$

Writing

$$W = \sum_i X_i, \qquad U_k = \sum_{i\neq k} X_i,$$

it follows that

$$\mathrm{Cov}(W) = \mathrm{Cov}(U_k) + \mathrm{Cov}(X_k).$$

Consequently, if $\mathrm{Cov}(W) = I_d$, and if P_k^2 denotes the covariance matrix of U_k while Σ_k denotes that of X_k, then

$$P_k^2 = I_d - \Sigma_k.$$

No individual covariance matrix is required to be nonsingular.

This is the covariance algebra in equation (3.2) and the paragraph immediately following it in Bentkus (2004), printed p. 403. The subsum identity above is stated at the level of covariance bilinear forms; the last display is its coordinate-matrix form and is exactly the paper's $P_k^2 = I - \operatorname{cov} X_k$ notation.

2.3 Whitening and covariance normalization

Theorem 2.3.1

Whitening a positive-definite covariance. Let S be a positive-definite real $d \times d$ matrix, let $R = S^{1/2}$ be its positive square root, and set $T = R^{-1} = S^{-1/2}$. Then

$$TST = I_d.$$

Equivalently, regarding these matrices as continuous linear maps,

$$T\big(S(Tx)\big) = x \qquad (x \in \mathbb{R}^d).$$

Bentkus (2004), defines the positive covariance square root C by $C^2 = \mathrm{Cov}(S_n)$ in the notation preceding Theorem 1.1, printed p. 400, and uses C^{-1} in the normalization following equation (3.1), printed p. 403. This gives the displayed $C^{-1}C^2C^{-1} = I$ identity.

Theorem 2.3.2

Gaussian pushforward under whitening. Under the preceding hypotheses, let $T = S^{-1/2}$ and let $R = S^{1/2}$. These maps are mutually inverse continuous linear equivalences, and their Gaussian pushforwards satisfy

$$T_{\#}N(0,S) = N(0,I_d), \qquad R_{\#}N(0,I_d) = N(0,S).$$

This is the Gaussian part of the rescaling $S_n \mapsto C^{-1}S_n$ and $Z \mapsto C^{-1}Z$ immediately after equation (3.1) in Bentkus (2004), printed p. 403; the covariance-square-root notation is introduced before Theorem 1.1, printed p. 400.

Theorem 2.3.3

The identity-covariance theorem contract. Let $X_1, \dots, X_n$ be independent, centered L^3 random vectors, suppose that their sum W has positive-definite covariance S, and put $\widehat{X}_i = S^{-1/2}X_i$. Then

$$\sum_i \widehat{X}_i = S^{-1/2}W,$$

the family $(\widehat{X}_i)_i$ remains independent and centered, and

$$\mathrm{Cov}\left(\sum_i \widehat{X}_i\right) = I_d.$$

For a fixed constant C, the identity-covariance contract asserts that every such normalized family and every measurable convex $B \subseteq \mathbb{R}^d$ satisfy

$$\left|\Pr\left(\sum_i \widehat{X}_i \in B\right) - \gamma_d(B)\right| \le Cd^{1/4}\sum_i \mathbb{E}\|\widehat{X}_i\|^3.$$

Here γ_d denotes standard Gaussian measure on $\mathbb{R}^d$.

This is the standardized estimate in equation (3.1) and the rescaling paragraph immediately after it in Bentkus (2004), printed p. 403. It separates the identity-covariance estimate from the subsequent whitening argument.

Theorem 2.3.4

One-set transport from the standardized estimate. Let S be positive definite, $T = S^{-1/2}$, $W = \sum_i X_i$, and $B = T(A)$. Write γ_S for centered Gaussian measure with covariance S. An invertible linear map preserves convexity, and the two event identities are

$$\Pr(TW \in B) = \Pr(W \in A), \qquad \gamma_{I_d}(B) = \gamma_S(A).$$

Consequently, any identity-covariance bound at the single set B transports to

$$|\Pr(W \in A) - \gamma_S(A)| \le C d^{1/4} \sum_i \mathbb{E}\|S^{-1/2} X_i\|^3.$$

Bentkus (2004), uses invariance of the set class under invertible symmetric linear maps in condition (i), printed p. 401, and applies the transformation C^{-1} immediately after equation (3.1), printed p. 403. This node is the one-set form of that reduction and yields exactly the normalization in Theorem 1.1, equation (1.1), printed p. 400.

2.4 Signed distance and coarea

Theorem 2.4.1

Signed distance and its eikonal identity. Let $A \subsetneq \mathbb{R}^d$ be a nonempty convex set and define

$$\delta_A(x) = d(x, A) - d(x, A^c) = \begin{cases} -d(x, A^c), & x \in A, \\ d(x, A), & x \notin A. \end{cases}$$

Then δ_A is Fréchet differentiable at Lebesgue-almost every point and

$$\|D\delta_A(x)\| = 1$$

at almost every such point.

Raič (2019), defines the signed distance on printed pp. 2825–2826 and proves the almost-everywhere differentiability and eikonal conclusions in Proposition 3.3(2)–(3), printed p. 2844. The derivative above is the Fréchet derivative on Euclidean space.

Lemma 2.4.2

Signed-distance fibers are parallel-set frontiers. Under the preceding hypotheses, for $t > 0$, $t = 0$, and $t < 0$, respectively,

$$\{x : \delta_A(x) = t\} = \partial\{x : d(x, \overline{A}) \le t\},$$
$$\{x : \delta_A(x) = 0\} = \partial A,$$
$$\{x : \delta_A(x) = t\} = \partial\{x \in A : B(x, -t) \subseteq A\}.$$

Here the outer parallel set is closed and the ball in the inner parallel set is open.

This is Proposition 3.3(4) of Raič (2019), printed p. 2844, with the positive and negative cases written separately to expose the exact outer-closure and open-ball conventions used by the setwise smoothing inequality.

Lemma 2.4.3

Outer and inner shells as signed-distance slabs. Let $A \subsetneq \mathbb{R}^d$ be nonempty and convex, let $\gamma_d = N(0, I_d)$, and let $\varepsilon \ge 0$. The convex frontier is Gaussian-null, and

$$\{x : d(x, \overline{A}) \le \varepsilon\} \setminus \overline{A} = \delta_A^{-1}((0, \varepsilon]),$$
$$A \setminus \{x \in A : B(x, \varepsilon) \subseteq A\} = A \cap \delta_A^{-1}((-\varepsilon, 0]).$$

Consequently the omitted boundary has no effect on mass and

$$\gamma_d\big(\{x : d(x, \overline{A}) \le \varepsilon\} \setminus A\big) = \gamma_d\big(\delta_A^{-1}((0, \varepsilon])\big),$$
$$\gamma_d\big(A \setminus \{x \in A : B(x, \varepsilon) \subseteq A\}\big) = \gamma_d\big(\delta_A^{-1}((-\varepsilon, 0])\big).$$

Raič (2019), introduces the signed-distance neighborhoods on printed pp. 2825–2826 and uses precisely these positive and negative slabs in the proof of Proposition 3.1, printed pp. 2844–2845, before applying the coarea formula. The identities above retain the equality boundaries; the convex frontier is Gaussian-null.

Theorem 2.4.4

Normalized affine codimension-one slicing. Put

$$\phi_d(x) = (2\pi)^{-d/2} e^{-\|x\|^2/2}.$$

Then $\gamma_d = \phi_d \, \lambda_d$ is standard Gaussian measure and $\int \phi_d \, d\lambda_d = 1$. If $v \ne 0$, $p \in \mathbb{R}^d$, and

$$H_t = p + tv + v^\perp,$$

then every measurable $B \subseteq \mathbb{R}^d$ and every nonnegative measurable w satisfy

$$\lambda_d(B) = \|v\| \int_{\mathbb{R}} \mathcal{H}^{d-1}(B \cap H_t) \, dt,$$
$$\int_{\mathbb{R}^d} w \, d\lambda_d = \|v\| \int_{\mathbb{R}} \int_{H_t} w \, d\mathcal{H}^{d-1} \, dt,$$
$$\gamma_d(B) = \|v\| \int_{\mathbb{R}} \int_{B \cap H_t} \phi_d \, d\mathcal{H}^{d-1} \, dt.$$

The Hausdorff measure is normalized so that top-dimensional measure is Euclidean volume. Moreover, if $u \in \mathbb{R}^{n+1}$ is a unit vector, then

$$\int_{au+u^{\perp}} \phi_{n+1}(y)\, d\mathcal{H}^n(y) = \phi_1(a) \le 1.$$

Ball (1993) uses codimension-one projections and Gaussian surface integrals in the proof of Theorem 4, printed pp. 415–419. Raič (2019) states the general area/coarea input as Proposition 3.2 and Corollary 3.1, printed pp. 2843–2844, and applies it in Proposition 3.1, printed pp. 2843–2845. This node supplies the Mathlib-normalized affine specialization; the nonlinear coarea formula for signed distance is recorded in the next node.

Theorem 2.4.5

Weighted area formula for an injective Lipschitz chart. Let U and V be finite-dimensional real inner-product spaces, put $m = \dim U$, and let $\varphi : S \subseteq U \to V$ be Lipschitz and one-to-one on a measurable set S. For every nonnegative measurable weight $g : V \to [0, \infty]$,

$$\int_{\varphi(S)} g(y)\, d\mathcal{H}^m(y) = \int_S J_m(D\varphi|_S)(x)\, g(\varphi(x))\, dx.$$

Here $D\varphi|_S$ is the within derivative and J_m is its absolute m-dimensional Jacobian. In particular, if a Lipschitz inverse chart φ has a linear left inverse P on S, so that $P(\varphi(x)) = x$, the same identity applies without separately assuming full rank: differentiating the composition makes $D\varphi|_S$ injective almost everywhere.

This is the injective weighted specialization of Corollary 3.2.32 in Federer (1969). Raič (2019) records the area formula as Proposition 3.2, printed pp. 2843–2844, and uses it both for coarea and for the boundary projection estimate.

Theorem 2.4.6

Weighted scalar coarea formula. Let $d \in \mathbb{N}$, let $f : \mathbb{R}^d \to \mathbb{R}$ be globally Lipschitz, and let $w : \mathbb{R}^d \to [0, \infty]$ be Borel measurable. Put

$$J_f(x) = \|Df(x)\|, \qquad Q_{f,w}(t) = \int_{f^{-1}(\{t\})} w(x)\, d\mathcal{H}^{d-1}(x),$$

where $Df(x)$ is taken to be zero at points where f is not differentiable and $d - 1$ is the natural-number predecessor. Then

$$\int_{\mathbb{R}^d} J_f(x) w(x)\, dx = \int_{\mathbb{R}} Q_{f,w}(t)\, dt.$$

For every $a, b \in \mathbb{R}$, the exact half-open slab form is

$$\int_{f^{-1}((a,b])} J_f(x) w(x)\, dx = \int_{(a,b]} Q_{f,w}(t)\, dt.$$

If $\|Df(x)\| = 1$ for Lebesgue-almost every x, this reduces to

$$\int_{f^{-1}((a,b])} w(x)\, dx = \int_{(a,b]} Q_{f,w}(t)\, dt.$$

Corollary 3.2.32 of Federer (1969) supplies the area/coarea theorem. Raič (2019) states the measurable-fiber theorem as Proposition 3.2, printed pp. 2843–2844, and its scalar coarea specialization as Corollary 3.1, printed p. 2844. Raič combines it with the almost-everywhere unit gradient of signed distance in Proposition 3.3 and the proof of Proposition 3.1, printed pp. 2844–2845. The version above uses nonnegative extended-valued weights, includes the exact $(a, b]$ convention, and also proves the dimension-zero endpoint under the displayed predecessor convention.

Theorem 2.4.7

Signed-distance coarea identities for Gaussian shell profiles. Let $A \subsetneq \mathbb{R}^d$ be nonempty and convex, put

$$\phi_d(x) = (2\pi)^{-d/2} e^{-\|x\|^2/2}, \qquad P_\gamma(C) = \int_{\partial C} \phi_d(x)\, d\mathcal{H}^{d-1}(x),$$

where $\mathcal{H}^{d-1}$ is normalized so that top-dimensional Hausdorff measure is Euclidean volume. For $t > 0$ and $t < 0$, respectively, define

$$p_A^+(t) = P_\gamma\big(\{x : d(x, \overline{A}) \le t\}\big), \qquad p_A^-(t) = P_\gamma(\{x : B(x, -t) \subseteq A\}),$$

with open balls in the inner parallel set. If δ_A is signed distance, then

$$p_A^+(t) = \int_{\{x:\delta_A(x)=t\}} \phi_d(x)\, d\mathcal{H}^{d-1}(x) \quad (t > 0),$$
$$p_A^-(t) = \int_{\{x:\delta_A(x)=t\}} \phi_d(x)\, d\mathcal{H}^{d-1}(x) \quad (t < 0).$$

Consequently, for every $\varepsilon \ge 0$, the exact half-open disintegrations are

$$\gamma_d\big(\{x : d(x, \overline{A}) \le \varepsilon\} \setminus A\big) = \int_{(0,\varepsilon]} p_A^+(t)\, dt,$$
$$\gamma_d(A \setminus \{x : B(x, \varepsilon) \subseteq A\}) = \int_{(-\varepsilon,0]} p_A^-(t)\, dt.$$

In particular, if every convex $C \subseteq \mathbb{R}^d$ satisfies $P_\gamma(C) \le K_d$ for $K_d = 4d^{1/4}$, then each of these two shell masses is at most $K_d\varepsilon$.

Raič (2019), Proposition 3.1 defines the boundary and shell suprema and derives both profile disintegrations, printed pp. 2843–2845. Corollary 3.1 gives the scalar coarea formula, and Proposition 3.3 identifies $\|\nabla \delta_A\| = 1$ almost everywhere and $\{\delta_A = t\} = \partial A^t$, both printed p. 2844. The two displayed outer and inner equalities occur in the proof of Proposition 3.1 on printed p. 2845. The node specializes Raič's continuous density to ϕ_d, retains the precise equality-boundary conventions, and records the direct perimeter-to-shell implication.

2.5 Ball's Gaussian perimeter theorem

Theorem 2.5.1

Intrinsic Hausdorff measure on the Euclidean unit sphere. Let $S^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$, let $\iota : S^{d-1} \hookrightarrow \mathbb{R}^d$ be inclusion, and let $\mathcal{H}_S^{d-1}$ and $\mathcal{H}_E^{d-1}$ denote Euclidean Hausdorff measure on the sphere as a metric space and on the ambient space, respectively. Then

$$\iota_{\#}\mathcal{H}_S^{d-1} = \mathcal{H}_E^{d-1} \restriction S^{d-1}.$$

Thus, for every nonnegative Borel function $F : \mathbb{R}^d \to [0, \infty]$,

$$\int_{S^{d-1}} F(\theta)\, d\mathcal{H}_S^{d-1}(\theta) = \int_{S^{d-1}} F(x)\, d\mathcal{H}_E^{d-1}(x).$$

If $d \geq 1$ and $\kappa_d = \lambda_d(B(0,1))$, then

$$\mathcal{H}_S^{d-1}(S^{d-1}) = d\kappa_d.$$

This measure is finite, is nonzero when $d \geq 1$, and every linear isometry $U : \mathbb{R}^d \to \mathbb{R}^d$ preserves it:

$$(U|_{S^{d-1}})_{\#}\mathcal{H}_S^{d-1} = \mathcal{H}_S^{d-1}.$$

Ball (1993), equation (1) and the centered-ball evaluation immediately following it, printed pp. 411–412, use codimension-one Hausdorff measure and the factor $r^{d-1} d\kappa_d$, which fixes the displayed surface-area normalization. Lemma 3 introduces rotation-invariant probability on the sphere, printed p. 414, and Theorem 4, equations (2)–(3), printed pp. 415–416, passes between spherical averaging and Hausdorff surface projection. The intrinsic-to-ambient pushforward identity makes that normalization explicit; the rotation statement is valid more generally for every Hausdorff dimension on any real inner-product space.

Theorem 2.5.2

Projection Jacobians and supporting-normal charts. Let $d \geq 2$ and $u, \theta \in S^{d-1}$. The orthogonal projection $P_{u,\theta} : u^{\perp} \to \theta^{\perp}$ satisfies

$$J_{d-1} P_{u,\theta} = |\langle \theta, u \rangle|.$$

If $C \subseteq \mathbb{R}^d$ is a compact convex body, every $x \in \partial C$ admits an outward supporting unit normal u, meaning

$$\langle y - x, u \rangle \leq 0 \quad \text{for every } y \in C.$$

For fixed θ, the boundary points with $\langle \theta, u \rangle \neq 0$ are the countable union of the two families

$$\langle \theta, u \rangle \geq a, \qquad \langle \theta, u \rangle \leq -a, \qquad a > 0.$$

On each such patch, projection onto $\theta^\perp$ is one-to-one and its inverse ψ is Lipschitz. At every differentiability point of ψ with unit normal u to its image tangent hyperplane,

$$|\langle \theta, u \rangle| \, J_{d-1} D\psi = 1.$$

This is the projection identity and graph decomposition in Ball (1993), equations (2)–(3) and the paragraph immediately following them, printed p. 416, in the proof of Theorem 4. The strict positive and negative patches make the inverse charts and their Jacobian cancellation explicit. Composing these charts with the preceding weighted area formula gives the surface-projection identity used in Ball's argument.

Theorem 2.5.3

Weighted Cauchy projection formula for the Euclidean sphere. Let $d \geq 2$, let $v \in S^{d-1}$, write $P_v : \mathbb{R}^d \to v^\perp$ for orthogonal projection, and let $g : v^\perp \to [0, \infty]$ be Borel measurable. With intrinsic codimension-one Hausdorff measure on S^{d-1} and Euclidean volume on $v^\perp$,

$$\int_{S^{d-1}} g(P_v \theta) |\langle \theta, v \rangle| \, d\mathcal{H}^{d-1}(\theta) = 2 \int_{B_{v^\perp}(0,1)} g(z) \, dz.$$

Equivalently, the same left-hand integral may be written over the unit sphere as an ambient subset with ambient $\mathcal{H}^{d-1}$. The factor 2 is exact: the two strict hemispheres each project one-to-one onto the open unit ball, while the equator has zero integrand.

Ball (1993), Theorem 4 constructs the projection measure in equation (2), then observes that almost every line in a fixed direction meets a convex boundary at most twice, printed pp. 415–416. The infinitesimal surface-projection factor $|\langle \theta, v \rangle|$ is equation (3), printed p. 416. Applied to the Euclidean ball, the two projection charts give the exact weighted Cauchy formula above, including its intrinsic-Hausdorff and ambient-restriction forms.

Theorem 2.5.4

Spherical projection average. Let σ_{d-1} be rotation-invariant probability measure on S^{d-1}. For $d \geq 1$ and $u \in \mathbb{R}^d$,

$$\int_{S^{d-1}} \langle \theta, u \rangle^2 \, d\sigma_{d-1}(\theta) = \frac{\|u\|^2}{d},$$

and hence

$$\int_{S^{d-1}} |\langle \theta, u \rangle| \, d\sigma_{d-1}(\theta) \leq \frac{\|u\|}{\sqrt{d}}.$$

The measure is invariant under every linear isometry of $\mathbb{R}^d$.

Ball (1993), averages the projection inequality over S^{d-1} immediately before equation (4), printed p. 416, in the proof of Theorem 4. The displayed second moment is the rotation-invariant covariance calculation underlying Ball's $d^{-1/2}$ factor; Cauchy–Schwarz gives the absolute first-moment inequality.

Theorem 2.5.5

Ball's spherical rearrangement inequality in the applied form. Let $d \geq 1$, let σ_{d-1} be rotation-invariant probability measure on S^{d-1}, and let $F : \mathbb{R} \to \mathbb{R}$ be continuous and nondecreasing on $[0, \infty)$. If $u, v \in S^{d-1}$ are orthogonal, $0 \leq \alpha \leq \pi/2$, and

$$w_\alpha = \cos(\alpha) u + \sin(\alpha) v,$$

then

$$\int_{S^{d-1}} F(|\langle \theta, u \rangle|) |\langle \theta, v \rangle| \, d\sigma_{d-1}(\theta) \leq \int_{S^{d-1}} F(|\langle \theta, u \rangle|) |\langle \theta, w_\alpha \rangle| \, d\sigma_{d-1}(\theta).$$

Ball (1993), Lemma 3 and its circle-fiber proof, printed pp. 414–415, treats two arbitrary nondecreasing functions on $[0, 1]$. The result recorded here is the exact continuous $G(t) = t$ specialization used in Theorem 4: on printed p. 416, immediately after equation (3), Ball takes $F(t) = f(r\sqrt{1 - t^2})$, which is nondecreasing because the radial density f is nonincreasing, and concludes that the spherical integral is minimized when the position and normal directions are perpendicular.

Lemma 2.5.6

Radial Gaussian peak and Gamma bound. For every integer $n \geq 3$ and every $t \geq 0$,

$$e^{-t^2/2} t^{n-2} \leq e^{-(n-2)/2} (n-2)^{(n-2)/2}.$$

Moreover,

$$\frac{e^{-(n-2)/2} (n-2)^{(n-2)/2}}{2^{n/2-1} \Gamma((n-1)/2) \sqrt{\pi}} \leq \frac{1}{\pi}.$$

These are the one-dimensional maximum and the normalized Gamma estimate in the last paragraph of Ball (1993), printed p. 419, completing the analytic estimate in the proof of Theorem 4. The second inequality retains Ball's exact $1/\pi$ constant in every dimension $n \geq 3$.

Theorem 2.5.7

Ball's radial projection majorant. For $n \geq 2$, put

$$a_n = (\sqrt{2\pi})^{-n}, \qquad h_n(t) = a_n \int_0^{\pi/2} (n{-}1{-}t^2 \sin^2\theta) \sin^{n-2}\theta \, e^{-t^2 \sin^2\theta/2} \, d\theta, \qquad f_n(t) = \max\{h_n(t), 0\}.$$

Then f_n is continuous on $\mathbb{R}$, nonnegative everywhere, nonincreasing on $[0, \infty)$, and

$$t \geq \sqrt{2n} \quad \Longrightarrow \quad f_n(t) = 0.$$

For a radial function $q : \mathbb{R} \to \mathbb{R}$, define its spherical projection by

$$(T_n q)(r) = \frac{2}{\pi} \int_0^{\pi/2} q(r \sin\theta) \sin^{n-1}\theta \, d\theta.$$

For every $r \in \mathbb{R}$, the signed auxiliary function has the exact projection identity and its positive part gives the Gaussian majorization

$$(T_n h_n)(r) = a_n e^{-r^2/2} \leq (T_n f_n)(r).$$

Finally, set

$$c_n = \left(2^{n/2-1} \Gamma\left(\frac{n-1}{2} \right) \sqrt{\pi} \right)^{-1}, \qquad p_n = (n{-}1) \, \mathrm{vol}_{n-1}(B_{n-1}(0,1)).$$

The polar normalization is exact, $p_n a_n = c_n$, and the one-dimensional and Euclidean mass bounds are

$$\frac{c_n}{a_n} \int_{(0,\infty)} f_n(t) t^{n-2} \, dt \leq 2n^{1/4}, \qquad \int_{\mathbb{R}^{n-1}} f_n(\|x\|) \, dx \leq 2n^{1/4}.$$

The same nonnegative-integral bound holds on every finite-dimensional real inner-product space of dimension $n - 1$.

Ball (1993), reduces Theorem 4 to the projection majorization in equation (4), defines the signed inverse transform by equations (5)–(6), and takes its positive part, printed p. 417. The power-series calculation following equation (7), also on printed p. 417, proves the required monotonicity; the zero-location estimate and polar mass identity are on printed p. 418. The two-part integral estimate and normalized Gamma peak close $\mu(\mathbb{R}^{n-1}) \leq 2n^{1/4}$ on printed pp. 418–419. The statement above separates the construction, the exact signed projection, the positive-part majorization, the compact positive-radius support, and both normalized forms of the mass estimate.

Theorem 2.5.8

Spherical majorization of Gaussian density times a normal magnitude. Let $d \geq 2$, let $x, N \in \mathbb{R}^d$, and write P_θ for orthogonal projection onto $\theta^\perp$. If f_d is Ball's radial majorant, ϕ_d is standard Gaussian density, and $\mathcal{H}_S^{d-1}$ is intrinsic Hausdorff measure on S^{d-1}, then

$$\phi_d(x)\|N\| \leq \frac{1}{\mathcal{H}_S^{d-1}(S^{d-1})} \int_{S^{d-1}} f_d(\|P_\theta x\|)|\langle \theta, N\rangle| \, d\mathcal{H}_S^{d-1}(\theta).$$

The statement also includes the exact orthogonal-direction identity with Ball's one-dimensional transform and the monotone spherical rearrangement inequality that extends it to an arbitrary normal direction. No normalization assumption is made on N.

Ball (1993), derives the spherical density inequality in equation (3), applies Lemma 3 to reduce arbitrary position and normal directions to the orthogonal case, and rewrites that case as equation (4), printed pp. 416–417. Equations (5)–(7) construct the monotone radial majorant used here, printed pp. 417–418. These steps give the pointwise vector inequality used in the boundary area formula.

Theorem 2.5.9

Weighted boundary projection area and the factor-two line multiplicity. Let $d \geq 2$, let $C \subseteq \mathbb{R}^d$ be a compact convex body with nonempty interior, and let $\theta \in S^{d-1}$. There is a finite disjoint family of Lipschitz graph charts $\varphi_j : T_j \to \partial C$. For every nonnegative measurable $w : \theta^\perp \to [0, \infty]$,

$$\sum_j \int_{T_j} J_{d-1}(P_\theta \circ D\varphi_j(z)) \, w(P_\theta \varphi_j(z)) \, dz \leq 2 \int_{\theta^\perp} w(y) \, dy.$$

The chart Jacobian is the normal-free representative of $|\langle \theta, n_C\rangle| \, d\mathcal{H}^{d-1}$. In particular, with Ball's radial majorant f_d,

$$\sum_j \int_{T_j} J_{d-1}(P_\theta \circ D\varphi_j(z)) \, f_d(\|P_\theta \varphi_j(z)\|) \, dz \leq 4d^{1/4}.$$

Ball (1993), introduces the projected comparison measure in equation (2), printed p. 415, and observes that almost every line in a fixed direction meets the boundary of a convex body at most twice, yielding the factor 2, printed p. 416 in the proof of Theorem 4. The radial mass estimate $\mu(\mathbb{R}^{d-1}) \leq 2d^{1/4}$ is completed on printed pp. 418–419. The same projection argument can be expressed through a finite disjoint chart cover and the equal-rank area formula, avoiding any measurable choice of an outward normal.

Theorem 2.5.10

Ball's Gaussian perimeter theorem. Let $d \geq 2$ and let $C \subseteq \mathbb{R}^d$ be convex. With

$$P_\gamma(C) = \int_{\partial C} (2\pi)^{-d/2} e^{-\|x\|^2/2}\, d\mathcal{H}^{d-1}(x), \qquad K_d = 4d^{1/4},$$

one has

$$P_\gamma(C) \leq K_d.$$

The conclusion holds without assuming that C is closed, bounded, or full-dimensional. For a compact convex body with nonempty interior it follows from the spherical density and boundary projection estimates above; closure, increasing compact truncation, affine slicing, and the empty-interior case complete the domain.

This is Ball (1993), Theorem 4, printed pp. 415–419, with its explicit constant $4d^{1/4}$. Ball states the theorem for convex bodies. The compact full-dimensional case is extended to every convex set by the boundary-content reductions described in the statement; the codimension-one case uses normalized affine slicing, while higher-codimension boundaries have zero $\mathcal{H}^{d-1}$ measure.

Theorem 2.5.11

Explicit Gaussian shell bounds for convex sets. Let $d \in \mathbb{N}$, let $A \subseteq \mathbb{R}^d$ be convex, and let $\varepsilon \geq 0$. Put $K_d = 4d^{1/4}$ and let γ_d be standard Gaussian measure. Then

$$\gamma_d\big(\{x : d(x, \overline{A}) \leq \varepsilon\} \setminus A\big) \leq K_d \varepsilon,$$

and, with open balls in the inner core,

$$\gamma_d(A \setminus \{x : B(x, \varepsilon) \subseteq A\}) \leq K_d \varepsilon.$$

These statements include dimensions zero and one as well as empty, full, unbounded, and lower-dimensional convex sets.

Ball (1993), Theorem 4, printed pp. 415–419, supplies the boundary-content constant in dimensions at least two. Raič (2019), Proposition 3.1 and its proof, printed pp. 2843–2845, derive the outer and inner shell integrals from Gaussian boundary content by signed-distance coarea. Bentkus (2004), uses the resulting $O(d^{1/4}\varepsilon)$ estimate as equation (1.4), printed p. 401. The version above keeps the exact half-open boundary conventions and includes the elementary zero- and one-dimensional endpoints.

2.6 Smoothing convex indicators

Lemma 2.6.1

Setwise Bentkus smoothing inequality. Let μ, ν be probability measures on $\mathbb{R}^d$, let A be measurable and convex, and let $\varepsilon > 0$. Define the closed outer parallel set and the inner core by

$$A_\varepsilon^+ = \{x : d(x, \overline{A}) \le \varepsilon\}, \qquad A_\varepsilon^- = \{x : B(x, \varepsilon) \subseteq A\},$$

where $B(x, \varepsilon)$ is the open ball. The inner core is closed, convex, and contained in A. There are total cutoffs φ_A and $\varphi_{A_\varepsilon^-}$ with values in $[0, 1]$, equal to 1 on their indexing set and equal to 0 respectively off A_ε^+ and off A. They are continuously differentiable and obey

$$\|D\varphi(x)\| \le \frac{2}{\varepsilon}, \qquad \|D\varphi(x) - D\varphi(y)\| \le \frac{8}{\varepsilon^2}\|x - y\|.$$

Moreover,

$$\begin{aligned} |\mu(A) - \nu(A)| \le\ & \max\left\{ \left| \int \varphi_A \, d\mu - \int \varphi_A \, d\nu \right|, \left| \int \varphi_{A_\varepsilon^-} \, d\mu - \int \varphi_{A_\varepsilon^-} \, d\nu \right| \right\} \\ & + \max\left\{ \nu(A_\varepsilon^+ \setminus A), \nu(A \setminus A_\varepsilon^-) \right\}. \end{aligned}$$

This is the setwise form of Lemma 2.1 in Bentkus (2004), combined with the cutoff of Lemma 2.2, equation (2.2), printed p. 402. The paper takes the outer neighborhood $\{d_A \le \varepsilon\}$ and defines its inner neighborhood using a closed ball. The inner core above uses an open ball, equivalently the complement of the open ε-thickening of A^c; this keeps the equality boundary in the displayed inner shell. Empty and nonclosed convex sets are handled by the total cutoff and $\overline{A}$ explicitly.

2.7 Gaussian density calculus

Lemma 2.7.1

Derivative integral against a Lipschitz function. Let E be a finite-dimensional real normed vector space with a Haar measure λ, let $p : E \to \mathbb{R}$ be a Schwartz function, and suppose that $f : E \to \mathbb{R}$ is a-Lipschitz. Then, for every $h \in E$,

$$\left| \int_E f(y) \, Dp(y)[h] \, d\lambda(y) \right| \le a\|h\| \int_{\operatorname{tsupp} f} |p(y)| \, d\lambda(y).$$

This is Bentkus (2004), Lemma 2.3, equations (2.3)–(2.4), printed pp. 402–403. The paper's hypothesis that p and all of its derivatives decay faster than every inverse power is represented by the Schwartz-class hypothesis above. The difference-quotient argument does not require the Lipschitz factor to be differentiable.

Theorem 2.7.2

Gaussian density derivatives and the cubic-contraction bound. Let E be a finite-dimensional real inner-product space, let $m = \dim E$, and set

$$\rho(x) = (2\pi)^{-m/2} e^{-\|x\|^2/2}.$$

Its first three directional Fréchet derivatives are

$$\begin{aligned} D\rho(x)[h] &= -\langle x, h\rangle \rho(x), \\ D^2\rho(x)[h,k] &= (\langle x, h\rangle\langle x, k\rangle - \langle h, k\rangle)\rho(x), \\ D^3\rho(x)[h,k,l] &= \big(-\langle x, h\rangle\langle x, k\rangle\langle x, l\rangle + \langle h, k\rangle\langle x, l\rangle + \langle h, l\rangle\langle x, k\rangle + \langle k, l\rangle\langle x, h\rangle\big)\rho(x). \end{aligned}$$

In particular,

$$D^3\rho(x)[w, w, g] = \rho(x) H_{w,g}(x),$$

where

$$H_{w,g}(x) = 2\langle w, g\rangle\langle x, w\rangle + \|w\|^2\langle x, g\rangle - \langle x, g\rangle\langle x, w\rangle^2.$$

The contraction $H_{w,g}$ is integrable under the standard Gaussian law γ_E. Writing $m_4 = \int_{\mathbb{R}} t^4 \, dN(0,1)(t)$, one has the dimension-free estimate

$$\int_E |H_{w,g}(x)| \, d\gamma_E(x) \le (3 + \sqrt{m_4}) \, \|w\|^2 \|g\|.$$

This supplies the exact Gaussian analytic input to the integration-by-parts estimate in Bentkus (2004), Lemma 2.3, equations (2.3)–(2.4), printed pp. 402–403. The twice-equal third derivative is equation (3.21), printed p. 405, and the absolute Gaussian integral estimate is the dimension-free content used in (3.22)–(3.23), printed pp. 405–406. The general Lipschitz integration-by-parts lemma itself remains the separate node immediately above.

Theorem 2.7.3

Gaussian integration by parts localized to the support of a Lipschitz factor. Put

$$\rho(y) = (2\pi)^{-d/2} e^{-\|y\|^2/2}.$$

The directional derivative $D\rho(\cdot)[w]$ is Lebesgue integrable. If $f : \mathbb{R}^d \to \mathbb{R}$ is globally Lipschitz and differentiable, then

$$\int_{\mathbb{R}^d} Df(y)[h]\rho(y) \, dy = -\int_{\mathbb{R}^d} f(y) D\rho(y)[h] \, dy.$$

Moreover, if f is L-Lipschitz, then, for all $w, h \in \mathbb{R}^d$,

$$\left| \int_{\mathbb{R}^d} f(y) D^2\rho(y)[w, h], dy \right| \le L\|h\| \int_{\operatorname{tsupp} f} |D\rho(y)[w]| \, dy.$$

This is Bentkus (2004), Lemma 2.3, equations (2.3)–(2.4), printed pp. 402–403. The first identity is the Gaussian-density integration-by-parts step used in equation (3.32), printed p. 407; the final estimate specializes Lemma 2.3 to $p = D\rho(\cdot)[w]$.

Theorem 2.7.4

Pulled-back cutoff derivative and its shell-localized bound. Let A be closed and convex, $\varepsilon > 0$, and let φ_A be the Bentkus cutoff. For $a, x \in \mathbb{R}^d$ and a linear map L, set

$$q(u) = D\varphi_A(a + Lu)[x].$$

Then q is $8\|L\|\|x\|/\varepsilon^2$-Lipschitz and

$$\operatorname{tsupp} q \subseteq \{u : a + Lu \in A_\varepsilon^+ \setminus \operatorname{int} A\}.$$

Consequently, for all $w, h \in \mathbb{R}^d$,

$$\left| \int q(u) D^2\rho(u)[w,h], du \right| \le \frac{8\|L\|\|x\|\|h\|}{\varepsilon^2} \cdot \int_{\{u : a+Lu \in A_\varepsilon^+ \setminus \operatorname{int} A\}} |D\rho(u)[w]| \, du.$$

This is the deterministic integration-by-parts step in Bentkus (2004), equation (3.32), printed p. 407, with the shell placement used in equations (3.30)–(3.34), printed pp. 407–408. The topological support retains the boundary of A, hence the exact shell removes only int A.

2.8 Taylor remainders and angle splitting

Lemma 2.8.1

First-order and two-shift Taylor remainders. Let E be a real normed vector space, let $f : E \to \mathbb{R}$ be continuously differentiable, and suppose that Df is L-Lipschitz. Then

$$|f(x+h) - f(x) - Df(x)[h]| \le \frac{L}{2}\|h\|^2.$$

For two shifts $v, w \in E$, freezing the derivative at x gives

$$|(f(x+v+w) - f(x+v)) - Df(x)[w]| \le L\|w\| \left(\|v\| + \frac{\|w\|}{2} \right).$$

This is the integral Taylor estimate used in Bentkus (2004), equation (3.28), printed p. 407, before the centering and covariance cancellations are applied. The first inequality keeps the exact $1/2$ obtained by integrating the linear derivative increment.

Theorem 2.8.2

Integrated second-order remainder for the first Gaussian-density derivative. Let ρ be the standard Gaussian density on $\mathbb{R}^d$, and let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be measurable with $|\varphi| \le 1$. For all $a, w, g \in \mathbb{R}^d$,

$$\left| \int \varphi(x) D^3\rho(x-a)[w,w,g], dx \right| \le (3 + \sqrt{m_4})\|w\|^2\|g\|,$$

where $m_4 = \int t^4 \, dN(0,1)(t)$. Consequently, for every $h, g \in \mathbb{R}^d$,

$$\left| \int \varphi(x) \big(D\rho(x-h)[g] - D\rho(x)[g] + D^2\rho(x)[h,g] \big) \, dx \right| \le \frac{3 + \sqrt{m_4}}{2} \|h\|^2\|g\|.$$

This is the cancellation-ready Taylor estimate in Bentkus (2004), equations (3.19)–(3.23), printed pp. 405–406. The factor $1/2$ is the integral of the second-order Taylor kernel, while the dimension-free constant is the absolute cubic Hermite contraction.

Lemma 2.8.3

Exact trigonometric integrals for Bentkus's angle split. If $0 < \gamma \le \pi/2$, then

$$\int_\gamma^{\pi/2} \frac{\cos\alpha}{\sin^2\alpha} \, d\alpha = \frac{1}{\sin\gamma} - 1 \le \frac{1}{\sin\gamma}.$$

Consequently, if $0 < \varepsilon \le 1$ and $\gamma = \arcsin\varepsilon$, then

$$\int_{\arcsin\varepsilon}^{\pi/2} \frac{\cos\alpha}{\sin^2\alpha} \, d\alpha = \varepsilon^{-1} - 1.$$

For $0 \le \varepsilon \le 1$, the complementary small-angle mass is exactly

$$\int_0^{\arcsin\varepsilon} \cos\alpha \, d\alpha = \varepsilon.$$

Bentkus (2004), splits the rotation at γ in equations (3.11)–(3.15), printed pp. 404–405. In the proof of (3.13), the integral $\int_0^\gamma \cos\alpha \, d\alpha$ occurs immediately after equation (3.33), printed p. 407. In the proof of (3.14), equations (3.39)–(3.40) are followed by the bound on $\int_\gamma^{\pi/2} \cos\alpha / \sin^2\alpha \, d\alpha$, printed p. 409. Bentkus chooses $\sin\gamma = \varepsilon$ after equation (3.8), printed p. 404; the node retains the exact antiderivatives before taking the inequalities used in the paper.

Lemma 2.8.4

Three-piece angle decomposition and coordinate summation. If F is interval-integrable on $[0,\gamma]$ and $[\gamma,b]$, and R is interval-integrable on $[\gamma,b]$, then

$$\int_0^b F = \int_0^\gamma F + \int_\gamma^b (F - R) + \int_\gamma^b R.$$

For $0 < \varepsilon \le 1$, if $K \ge 0$ and

$$|F(\alpha)| \le K \frac{\cos \alpha}{\sin^2 \alpha} \quad \text{for } \alpha \in (\arcsin \varepsilon, \pi/2],$$

then

$$\left| \int_{\arcsin \varepsilon}^{\pi/2} F(\alpha)\, d\alpha \right| \le \frac{K}{\varepsilon}, \qquad \frac{\pi}{2} - \arcsin \varepsilon \le 2.$$

Suppose a coordinate contribution satisfies $T = S + L + R$ and, for nonnegative $d_{1/4}, q, \beta_k, K_L, K_R$,

$$|S| \le K_S d_{1/4}(1+q)\beta_k, \quad |L| \le K_L d_{1/4} q \beta_k, \quad |R| \le K_R d_{1/4} \beta_k.$$

Then

$$|T| \le (K_S + K_L + K_R) d_{1/4}(1+q)\beta_k.$$

Moreover, if $\beta = \sum_k \beta_k$, $\varepsilon > 0$, and

$$|T_k| \le K d_{1/4} \left(1 + \frac{C\beta}{\varepsilon}\right) \beta_k,$$

then

$$\left| \sum_k T_k \right| \le K d_{1/4} \left(\beta + \frac{C\beta^2}{\varepsilon}\right).$$

Deleting one nonnegative coordinate cannot increase the corresponding finite sum:

$$\sum_{i \ne k} \beta_i \le \sum_i \beta_i.$$

This is the scalar assembly behind Bentkus (2004), Section 3, printed p. 404. The small-angle, large-angle difference, and Gaussian-reference estimates are equations (3.13)–(3.15); they combine to equation (3.6), summation over coordinates gives equation (3.7), and insertion into the smoothing inequality gives equations (3.8)–(3.9). The displayed decomposition and omitted-coordinate inequality make explicit the bookkeeping compressed in the paper.

Lemma 2.8.5

Closure of the smoothing parameter. Let $C \ge 1$, $d \ge 1$, $\beta > 0$, and $\Delta \le 1$. Suppose

$$K(2\sqrt{C} + 1) \le C$$

and, whenever $\beta\sqrt{C} < 1$,

$$\Delta \le K d^{1/4} \left(\beta\sqrt{C} + \beta + \frac{C\beta^2}{\beta\sqrt{C}}\right).$$

Then

$$\Delta \le C d^{1/4} \beta.$$

This is the scalar calculation in Bentkus (2004), equations (3.9)–(3.10), printed p. 404. It chooses $\varepsilon = \beta\sqrt{C}$; the complementary branch $\varepsilon \geq 1$ closes directly from $\Delta \leq 1$.

2.9 The identity-covariance replacement induction

Lemma 2.9.1

The small-cardinality and large-summand induction branches. Let $X_1, \dots, X_n$ be independent centered L^3 random vectors in $\mathbb{R}^d$, assume that $\mathrm{Cov}(\sum_i X_i) = I_d$, and put

$$\beta = \sum_i \mathbb{E}\|X_i\|^3.$$

Then

$$\sum_i \mathbb{E}\|X_i\|^2 = d, \qquad d^3 \leq n\beta^2.$$

If $d > 0$, $M \geq 0$, and $n \leq d^3 M^2$, then $1 \leq M\beta$; hence, for every probability measure ν on $\mathbb{R}^d$ and every set A,

$$\left| \Pr\left(\sum_i X_i \in A \right) - \nu(A) \right| \leq M\beta.$$

Separately, if one summand satisfies $\mathbb{E}\|X_k\|^2 \geq \frac{1}{4}$, then

$$1 \leq 8\,\mathbb{E}\|X_k\|^3.$$

These are the two trivial branches at the start of Section 3 in Bentkus (2004), printed p. 403. The first is the Hölder calculation and $\Delta \leq 1$ argument immediately after equation (3.1) for $n \leq d^3 M^2$; the second is the large-individual-covariance exclusion immediately before equation (3.2). These are precisely the moment consequences used before the nontrivial Taylor induction begins.

Lemma 2.9.2

Leave-one-out whitening in the small-second-moment branch. Under identity total covariance, fix k and suppose that

$$\mathbb{E}\|X_k\|^2 < \frac{1}{4}.$$

If $U_k = \sum_{i \neq k} X_i$ and $S_k = \mathrm{Cov}(U_k)$, then S_k is positive definite. Its inverse square root $T_k = S_k^{-1/2}$ satisfies

$$\|T_k x\| \leq 2\|x\|, \qquad \|T_k x\|^3 \leq 8\|x\|^3.$$

Consequently, for every L^3 random vector Z on any measure space,

$$\int \|T_k Z\|^3 \leq 8 \int \|Z\|^3.$$

This is the small-individual-covariance branch introduced after equation (3.2) in Bentkus (2004), printed pp. 403–404. The paper writes $P_k^2 = \mathrm{Cov}(U_k)$ and $Q_k = P_k^{-1}$ and, after excluding the large-covariance trivial case, assumes $\|Q_k\| \le 2$; this gives the third-moment factor $2^3 = 8$.

Theorem 2.9.3

> **Identity-covariance convex-set replacement bound.** There is an absolute constant $C > 0$ such that, for every $d, n \in \mathbb{N}$ with $d > 0$, every probability space $(\Omega, \mathcal{F}, \mu)$, and every family $X_i : \Omega \to \mathbb{R}^d$ indexed by $i \in \mathrm{Fin}(n)$ satisfying $X_i \in L^3(\mu)$, mutual independence, and $\int_\Omega X_i \, d\mu = 0$, if $W = \sum_i X_i$ has identity covariance in the bilinear sense
>
> $$\mathrm{Cov}_{\mu \circ W^{-1}}(x, y) = \langle x, y \rangle \qquad \text{for every } x, y \in \mathbb{R}^d,$$
>
> one has, for every measurable convex $A \subseteq \mathbb{R}^d$,
>
> $$\left|(\mu \circ W^{-1})(A) - \gamma_d(A)\right| \le C d^{1/4} \sum_{i \in \mathrm{Fin}(n)} \int_\Omega \|X_i(\omega)\|^3 \, d\mu(\omega),$$
>
> where γ_d is standard Gaussian measure on $\mathbb{R}^d$.

This is the identity-covariance specialization of Bentkus (2004), Theorem 1.2, printed p. 401, and the standardized setup in equation (3.1), printed p. 403. The leave-one-out induction begins with equation (3.2), printed p. 403; the rotation and replacement estimates in equations (3.4)–(3.15) occupy printed pp. 404–405, and the proof continues through printed p. 409. Bentkus leaves the absolute constant unspecified; tracking constants through the preceding estimates gives one absolute choice.

2.10 Bentkus's convex-set theorem

Theorem 2.10.1

> **Bentkus's Lyapunov bound for convex sets.** There is an absolute constant $C > 0$ with the following property. Let $d, n \in \mathbb{N}$ with $d > 0$, let $X_1, \dots, X_n$ be independent measurable random vectors in $\mathbb{R}^d$ such that
>
> $$\mathbb{E} X_i = 0, \qquad \mathbb{E}\|X_i\|^3 < \infty,$$
>
> and set $W = \sum_{i=1}^n X_i$. Suppose that the covariance matrix $S = \mathrm{Cov}(W)$ is positive definite, and let G be a centered Gaussian random vector with covariance matrix S. Then for every measurable convex set $A \subseteq \mathbb{R}^d$,
>
> $$|\Pr(W \in A) - \Pr(G \in A)| \le C d^{1/4} \sum_{i=1}^n \mathbb{E}\left\| S^{-1/2} X_i \right\|^3.$$

This is Theorem 1.1 and equation (1.1) of Bentkus (2004), printed pp. 400–401. The matrix S here is the total covariance; $S^{-1/2}$ is the paper's inverse of the positive covariance square root. The proof first establishes the standardized replacement estimate and then applies covariance-square-root whitening.

References

Keith Ball (1993). "The reverse isoperimetric problem for Gaussian measure". *Discrete and Computational Geometry.* ***10****(4), pp. 411–420.*

Vidmantas Bentkus (2003). "On the dependence of the Berry–Esseen bound on dimension". Journal of Statistical Planning and Inference. **113**(2), pp. 385–402.

Vidmantas Bentkus (2004). "A Lyapunov-type bound in R^d". *Teoriya Veroyatnostei i ee Primeneniya.* ***49****(2), pp. 400–410.*

A. Bikelis (1966). "On estimates of the remainder term in the central limit theorem". Lithuanian Mathematical Journal. **6**(3), pp. 323–346.

Louis H. Y. Chen and Qi-Man Shao (2001). "A non-uniform Berry–Esseen bound via Stein's method". *Probability Theory and Related Fields.* ***120****(2), pp. 236–254.*

Louis H. Y. Chen and Qi-Man Shao, 2005. "Stein's method for normal approximation". In An Introduction to Stein's Method, pp. 1–59. (Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, vol. 4)

Herbert Federer, 1969. "Geometric Measure Theory". In *Springer. (Grundlehren der mathematischen Wissenschaften, vol. 153)*

Martin Raič (2019). "A multivariate Berry–Esseen theorem with explicit constants". Bernoulli. **25**(4A), pp. 2824–2853.